شریف، عاطفه؛ داورپناه، محمدرضا؛ صنعت‌جو، اعظم(۱۳۹۳). تحلیل و مقایسه فهرست سرعنوان‌های موضوعی فارسی با فهرست سرعنوان‌های موضوعی کنگره از منظر ویژگی‌های ساختار زبانی و شبکه‌معنایی. *پژوهش‌نامه کتابداری واطلاع رسانی*، ۴ (۱)، ۷۲- ۴۷.

تحلیل و مقایسه فهرست سرعنوان‌های موضوعی فارسی با فهرست سرعنوان‌های موضوعی کنگره از منظر ویژگی‌های ساختار زبانی و شبکه‌معنایی[۱]

عاطفه شریف [۲]، دکتر محمدرضا داورپناه [۳]، دکتر اعظم صنعت‌جو [۴]

تاریخ دریافت: ۹۱/۳/۲۴ تاریخ پذیرش: ۹۲/۳/۲۸

چکیده

هدف: *فهرست سرعنوان‌های موضوعی فارسی به منزله ابزاری استاندارد و ملی جهت بازنمایی موضوعی کتاب‌های فارسی از سوی کتابخانه ملی ایران و سایر کتابخانه‌ها در ایران مورد استفاده قرار می‌گیرد. هدف از انجام این پژوهش، بازشناسی عمیق ویژگی‌های سرعنوان‌های موضوعی فارسی در دو حیطه ساختار زبانی و شبکه معنایی است.*

روش: *پژوهش از نوع کاربردی است و به روش تحلیل محتوا انجام پذیرفته است. دستور کار کدگذاری مطابق با پرسش‌های پژوهش، چارچوب تعریف‌شده، مبانی نظری به ویژه در حوزه زبان و بازنگری متخصصان آن حوزه تدوین شد. فهرست سرعنوان‌های موضوعی فارسی ویرایش سوم و دو پیوست آن (۱۱۲۳۳ رکورد) و سرعنوان‌های موضوعی مشترک در فهرست کنگره و فارسی (۹۵۱۹ رکورد) به منزله جامعه آماری قرار گرفت. بخشی از داده‌ها از سطح جامعه آماری و بخشی دیگر از طریق نمونه تصادفی گردآوری شد.*

یافته‌ها: *طبق یافته‌ها از نظر ساختار زبانی، سرعنوان‌های موضوعی فارسی با ۸۶/۱۷ درصد از زبان طبیعی پیروی کرده‌اند و ۷۹/۸۷ درصد از سرعنوان‌ها پیش‌همارای طبیعی هستند. از نظر دستوری شکل غالب سرعنوان‌ها به صورت اسم (۹۶/۱ درصد) و گروه اسمی (۸۷ درصد) است. تفاوت معناداری میان دو فهرست از نظر طبیعی یا مصنوعی بودن زبان، وجود یادداشت دامنه و وجود شماره رده برقرار است. از نظر شبکه معنایی، تفاوتی اساسی و معنادار میان دو فهرست سرعنوانی موضوعی فارسی و*

---

۱. این مقاله برگرفته از پایان نامه دوره دکتری نویسنده نخست به راهنمایی دکتر محمدرضا داورپناه و مشاوره دکتر اعظم صنعت‌جو است.

۲. دانشجوی دوره دکترای علم اطلاعات و دانش شناسی دانشگاه فردوسی مشهد ، atefehsharif@gmail.com

۳. استاد گروه علم اطلاعات و دانش شناسی دانشگاه فردوسی مشهد ، mrdavarpanah@yahoo.com

۴. استادیار گروه علم اطلاعات و دانش شناسی دانشگاه فردوسی مشهد ، sanatjoo@gmail.com

کنگره قابل مشاهده است. با توجه به ضعف و اشکال در شبکه روابط معنایی فهرست سرعنوان‌های موضوعی فارسی لازم است این فهرست در جهت اصلاح و غنی‌سازی شبکه معنایی میان سرعنوان‌ها مورد بازنگری قرار گیرد.

کلیدواژه‌ها: *سرعنوان موضوعی فارسی؛ ساختار زبانی؛ شبکه معنایی؛ سرعنوان موضوعی کنگره.*

مقدمه و بیان مسئله

فهرست سرعنوان‌های موضوعی فارسی به منزله ابزاری استاندارد و ملی جهت بازنمایی موضوعی کتاب‌های فارسی از سوی کتابخانه ملی ایران و سایر کتابخانه‌ها در ایران مورد استفاده قرار می‌گیرد. سرعنوان‌های موضوعی فارسی از نظر ساختار از سرعنوان‌های موضوعی کتابخانه کنگره پیروی کرده است، با این وجود، به دلایل پشتوانه انتشاراتی، تنوع موضوع‌ها و نیز تفاوت‌های زبانی، دارای ویژگی‌های خاص خود نیز می‌باشد. انتظار می‌رود که واژگان موضوعی در سرعنوان‌های موضوعی فارسی به موازات تولد در بستر نگارش‌های علمی و ادبی– در این مورد به طور اختصاصی کتاب‌ها- گسترش یابد و بازنمونی از مفاهیم حوزه‌های موضوعی موجود در زبان فارسی باشد.

سرعنوان‌های موضوعی فارسی نشان‌گر ویژگی‌های زبان فارسی در بازنمون نظام موضوعی است و از سویی می‌تواند[۱] نشانی از کاربرد واژگان موضوعی در دنیای واقعی علم و ادب و نزد عالمان و نویسندگان حوزه‌های موضوعی گوناگون باشد.

سرعنوان‌های موضوعی، نه تنها به دنیای سنتی کتابخانه‌ها محدود نمی‌شود، به جای آن، همان‌گونه که یی و چان[۲] بیان داشته‌اند، به منزله ابزار کنترل منابع دیجیتالی نیز کاربرد یافته است (Yi, Chan, 2010). تاکنون سرعنوان‌های موضوعی کنگره در انجام طرح‌ها[۳] و تولید[۴] برخی نرم‌افزارهای کاربردی، مورد استفاده قرار گرفته است؛ از جمله اختصاص خودکار سرعنوان موضوعی کنگره به رکوردها در منابع الکترونیک (Frank, Paynter, 2004)، بهره‌گیری از سرعنوان‌های موضوعی کنگره در بخش کاوش کتاب موتور کاوش گوگل با هدف بهبود و سهولت در بازیابی مجموعه‌ای از آثار هم موضوع (Steve,2008) امکان‌سنجی و پیشنهاد بهره‌گیری از سرعنوان‌های موضوعی کنگره در فرایند تعریف

---

[۱]. چنان‌چه گزینش سرعنوان‌های موضوعی بر پایه کاربرد اصطلاحات تخصصی در حوزه‌ها باشد.

[۲] Yi & Chan

[۳] Projects

[۴] Develop

برچسب‌های اجتماعی[1] در فولکسونومی‌ها (Yi, Chan, 2010; Yi, Chan, 2009; Chan, 2009)[2]در محیط وب ۲ [3] و حتی کدگذاری سرعنوان‌های موضوعی کنگره با استاندارد نظام ساده سازماندهی دانش (اسکاس)[4] به طوری که به عنوان داده‌های پیوندی[5] در وب ۳[6] کاربرد دارد (Harper, 2006; Ed, Antoine, Clay, Dan, 2008). به این ترتیب واضح است که سرعنوان‌های موضوعی به منزله ابزار کنترل واژگان، اگر چه در دنیای سنتی کتابخانه‌ها متولد شده است اما حیات خویش را تا نظام‌های ذخیره و بازیابی پیوسته – وب– وب ۲ و حتی وب ۳ ادامه داده است. می‌توان علت این ماندگاری را، قابلیت ویژه‌ی سرعنوان‌های موضوعی در سازماندهی اطلاعات دانست.

پرسش‌های پژوهش

بررسی‌های مقدماتی پژوهشگر نشان از آن دارد که تاکنون پژوهش‌های اندکی فهرست سرعنوان‌های موضوعی فارسی را مورد توجه قرار داده‌اند و از آن میان، اندک پژوهشگری در شناسایی و تحلیل ویژگی‌های زبانی و شبکه معنایی فهرست سرعنوان‌های موضوعی فارسی به کنکاشی عمیق پرداخته است. به این ترتیب در مورد ویژگی‌های زبانی سرعنوان‌های موضوعی مشخص نیست که:

۱- ویژگی‌های ساختار زبانی[7] سرعنوان‌های موضوعی فارسی[1] از نظر تعداد و نسبت سرعنوان‌هایی با ساختار واژه به سرعنوان‌هایی با ساختار گروه چگونه است؟

---

۱ Social tags

۲ folksonomies

۳ Web 2.0

۱. SKOS: Simple Knowledge Organization System .یکی از استانداردهای مطرح شده از سوی کنسرسیوم وب جهانی می‌باشد و در کدگذاری نظام‌های دانش مانند طرح‌های رده‌بندی، سرعنوان‌های موضوعی و اصطلاح‌نامه‌ها کاربرد دارد. این استاندارد زبانی به منزله یکی از استانداردهای وب معنایی در کدگذاری هستی‌شناسی‌ها نیز شناخته می‌شود (Miles, Perez-Aguera, 2007, p 70)

۲. Linked data .حاصل این طرح در سایت رسمی کتابخانه کنگره (http://id.loc.gov) منتشر شده است.

۶ Web 3.0

۷. در تحلیل ساختار زبانی تعیین مقوله‌های دستوری (مقوله واژگانی و گروهی) سرعنوان‌های موضوعی و برخی از ویژگی‌های زبانی آن‌ها از جمله ساختمان واژه (ساده، مرکب، مشتق، مشتق –مرکب)، جمع و مفرد، تعداد واژه‌ها و طبیعی و مصنوعی بودن آن‌ها مورد توجه است. هر سرعنوان موضوعی می‌تواند در یکی از دو مقوله دستوری واژگانی یا گروهی قرار گیرد. سرعنوان‌هایی که از یک واژه ساخته شده‌اند در مقوله دستوری واژگانی قرار می‌گیرند و یکی از انواع اسم، صفت، فعل، قید و

۱-۱ انواع و نسبت واژه‌های مورد استفاده در سرعنوان‌های موضوعی فارسی از نظر مقوله دستوری واژگانی[۲]- اسم، صفت، قید، فعل و حرف - و ویژگی ساختمان واژه‌ها در آن مقوله – ساده، مشتق، مرکب و مشتق – مرکب- کدام است؟

۱-۲ چه تنوع و نسبتی از واحد زبانی گروه[۳] در فهرست سرعنوان‌های موضوعی فارسی قابل شناسایی است؟

۲- نسبت ساختار زبان طبیعی به زبان مصنوعی - سرعنوان‌های مقلوب[۴] و سرعنوان به همراه تقسیم فرعی- در سرعنوان‌های موضوعی فارسی چگونه است؟

۳- سطح پیش‌همارایی و نسبت سرعنوان‌های پیش‌همارا با ساخت طبیعی (پیش‌همارایی طبیعی) به سرعنوان‌های پیش‌همارا[۵] با ساخت مصنوعی (پیش‌همارایی مصنوعی) در فهرست سرعنوان‌های موضوعی فارسی چگونه است؟

پیرامون شبکه معنایی[۶] سرعنوان‌های موضوعی و پیوند میان موضوعات نیز مشخص نیست که:

۴- شبکه معنایی حاکم بر سرعنوان‌های موضوعی فارسی به لحاظ پیوند میان سرعنوان‌ها چه ویژگی‌هایی دارد؟

۴-۱ ساختار شبکه جزئی روابط معنایی[۷](LRS) در فهرست سرعنوان‌های موضوعی فارسی به لحاظ تعداد[۸] و شمار[۹] انواع روابط معنایی حاکم بر آن چگونه است؟

۴-۲ نسبت پیوستگی[۱] و دو مکمل آن یعنی نسبت وابستگی[۲] و انسجام[۳] میان سرعنوان‌ها در فهرست سرعنوان‌های موضوعی فارسی چگونه است؟

حرف خواهند بود. سرعنوان‌هایی با ساختار واژه از نظر ساختمان واژه نیز حالت‌های متفاوتی خواهند داشت. سرعنوان‌هایی که از بیش از یک واژه ساخته شده‌اند در مقوله دستوری گروه قرار می‌گیرند و و متناسب با نوع هسته‌ی گروه، انواعی دارند.

۱. سرعنوان‌های انتخاب شده

۲. منظور از مقوله‌های دستوری، پنج مقوله‌ی اسم، صفت، قید، فعل و حرف می‌باشد.

۳. منظور از گروه‌ها در این پژوهش۱.گروه‌های اسمی شامل گروه وصفی، اضافی، عطفی و مقید؛ ۲. گروه صفتی؛ ۳. گروه قیدی؛ ۴. گروه فعلی؛ ۵. گروه حرف اضافه‌دار و ۶. ترکیبی از گروه‌ها می‌باشد.

۴ Inverted headings

۵ pre-coordinaton level

۶ Semantic network

۷ Local Relational Structure

۸ Sum

۹ Count

۵- الگوی شبکه معنایی حاکم بر سرعنوان‌های موضوعی فارسی در مقایسه با سرعنوان‌های موضوعی کنگره چگونه است؟

۵-۱ ساختار شبکه جزئی (LRS) روابط معنایی از نظر تعداد و شمار روابط معنایی در سرعنوان‌های موضوعی فارسی در مقایسه با سرعنوان های موضوعی کنگره چگونه است؟

۵-۲ نسبت پیوستگی و دو مکمل آن یعنی نسبت وابستگی و انسجام سرعنوان‌های موضوعی فارسی در مقایسه با سرعنوان‌های موضوعی کنگره چگونه است؟

در تدوین چارچوب پژوهش، دو پیش‌فرض از سوی پژوهش‌گران پذیرفته شده است:

- از آن جایی که واژگان موضوعی در فهرست سرعنوان‌های موضوعی به مرور زمان و به واسطه پشتوانه انتشاراتی به زبان فارسی تولید می‌شوند و بازنمونی از مفاهیم حوزه‌های موضوعی به زبان فارسی تلقی می‌گردند، منطقی به نظر می‌رسد که سرعنوان‌های موضوعی فارسی نشانگر ویژگی‌های ساختاری زبان فارسی در بازنمون نظام موضوعی باشد.
- اگر چه از ابتدای شکل‌گیری فهرست سرعنوان‌های موضوعی کنگره و به ویژه بعد از سال ۱۹۷۵ به واسطه عدم رعایت منطقی روابط اصطلاح‌نامه‌ای و تبدیل ماشینی روابط اولیه به روابط اصطلاح‌نامه‌ای - هم‌ارز، سلسله‌مراتب و هم‌بسته - این ابزار مورد انتقاد قرار گرفت، این فهرست، الگوی توسعه سرعنوان‌های موضوعی فارسی قرار گرفته است. به این ترتیب به نظر می‌رسد که در شرایط فعلی، نوعی پذیرش و توافق ضمنی در مورد روابط معنایی تعریف شده در این ابزار، به وجود آمده است.

مروری بر پیشینه پژوهش

در نگاهی جامع بر پیشینه پژوهش‌های انجام گرفته در ایران مشخص می‌گردد که اندک پژوهش‌های انجام شده پیرامون فهرست سرعنوان‌های موضوعی فارسی بر سنجش تطابق و هم‌خوانی واژگانی (Bozorgi, 1991; Fattahi, Arastoopoor, 2007) و رعایت یکدستی و قانون‌مندی در ساخت سرعنوان‌های موضوعی فارسی و مقایسه آن با سرعنوان‌های موضوعی کنگره (Zhoadi, 2010) استوار است. سایر پژوهش‌های انجام گرفته در ایران بر اصطلاح‌نامه‌ها صورت پذیرفته است. از جمله ارزیابی جنبه‌های کیفی (Kazerani, Asadi gorgani & Rahadoost, 2002) و کمی (AmirHosseini, 2005, 2007a, 2007b, 2007c) اصطلاح‌نامه‌ها، بررسی تطبیقی چند اصطلاح‌نامه با یکدیگر از نظر ساختار

[1] Connectedness ratio
[2] Relativeness Ratio
[3] Integration Ratio

توصیفگرها (Akhondi, Horri, 2006; Khani, Alibeyg, Haghani, 2007) و شبکه روابط سلسله مراتبی (Khojoui, 1993; Mirzabeigi, 2005; Tabib, 2001) و مواردی چون تأثیر ورود اصطلاحات جدید بر ساختار روابط معنایی اصطلاح‌نامه‌ها (Mirzabeigi, Horri, 2006).

اخوتی، داورپناه، دیانی و شعبانی ورکی (Okhovati, DAvarpanah, Dayani, & Shabani, 2009) در پژوهشی به بررسی وضعیت تبعیت یا عدم تبعیت سرعنوان‌های موضوعی فارسی از واژه‌ها و عبارات موضوعی متون کتابداری و اطلاع‌رسانی بر اساس رویکرد تحلیل حوزه پرداختند. یافته‌ها نشانگر آن است که طبقه مفهومی منابع اطلاعاتی در متون مورد بررسی بیش از سایر طبقات مفهومی مورد توجه قرار گرفته است. شبکه مفهومی حاکم بر کتاب‌ها با شبکه مفهومی مقالات تفاوت معناداری داشته است و تنها ۱۰/۶۵ درصد از مفهوم- واژگان متون مورد بررسی در بین مفاهیم مرتبط با کتابداری و اطلاع‌رسانی سرعنوان‌های موضوعی فارسی و ۲۹/۳۲ درصد آن‌ها در کل سرعنوان‌های موضوعی فارسی وجود داشته‌اند. بر این اساس تفاوت معناداری میان شبکه‌های مفهومی سرعنوان‌های موضوعی فارسی و متون مورد بررسی مشاهده شده است. وی همچنین گزارش کرده است که شبکه مفهومی سرعنوان‌های موضوعی فارسی به طور معناداری با شبکه مفهومی سرعنوان موضوعی کنگره، سرعنوان موضوعی سیرز و لیزا متفاوت بوده است.

پیشینه پژوهش در بخش خارج از کشور نشان از تعداد بیشتر پژوهش‌ها درباره ویژگی‌های سرعنوان‌های موضوعی کنگره دارد. به طوری که یی و چان (Yi, Chan, 2010) در پژوهشی عمیق بر ویژگی‌های سرعنوان‌های موضوعی کنگره متمرکزشده‌اند و ویژگی‌های ساختار زبانی و معنایی سرعنوان‌ها را تعیین نموده‌اند. این پژوهش‌گران باهدف امکان‌سنجی کاربرد سرعنوان‌های موضوعی کنگره در فضای شبکه‌های دیجیتالی، به بازبینی و بررسی ویژگی‌های نحوی[1] و ساختاری[2] سرعنوان‌های موضوعی (۳۴۱۷۴۵ رکورد) پرداخته‌اند. به منظور سهولت در تحلیل داده‌ها، چند برنامه با زبان Perl نیز نوشته و مورد استفاده قرار گرفته است.

در پاسخ به نخستین پرسش، دو نرم‌افزار برچسب‌زن[3] به زبان انگلیسی انتخاب شد تا نقش کلمه‌ها[4] را به صورت خودکار برچسب‌گذاری کند. نتایج حاصل از تحلیل‌ها نشان داد که حدود ۶۷/۷ درصد از

---

۱ Syntactical

۲ Structural

۳. Tagger. در این پژوهش Monty tagger (version 1.2) و TreeTagger مورد استفاده قرار گرفته است

۴ Parts of Speech (POS)

سرعنوان‌ها در فهرست سرعنوان‌های موضوعی کنگره، بدون تقسیم فرعی هستند و ۳۲/۳ درصد از سرعنوان‌ها به همراه تقسیمات فرعی ظاهرشده‌اند. سرعنوان‌های تک کلمه‌ای یا ساده ۱۲/۶ درصد هستند و سرعنوان‌های دو کلمه‌ای بیش‌ترین رخداد یعنی ۳۶/۹ درصد را به خود اختصاص داده‌اند. در این مورد با افزایش تعداد کلمات، رخداد سرعنوان‌ها سیر نزولی می‌یابد. سرعنوان‌های دارای توضیحگر[۱] الگوی حاکم بر سرعنوان‌های مرکب بوده است. پس از آن به ترتیب سرعنوان‌های مقلوب[۲]، گروه صفتی[۳]، گروه حرف اضافه‌دار[۴]، گروه عطفی[۵] و در نهایت سرعنوان‌های شناور[۶] آمده‌اند. پژوهشگران ضمن تأکید بر اهمیت پیش‌بینی‌پذیر بودن[۷] الگوهای حاکم بر سرعنوان‌های موضوعی در راستای گزینش و تخصیص خودکار سرعنوان به منابع اطلاعاتی، به این نتیجه رسیده‌اند که بیش از ۹۰ درصد از سرعنوان‌ها شامل اشکال متنوعی از گروه‌ها از جمله تقسیم فرعی، توضیحگرها، سرعنوان‌های شناور و ... در بازنمایی موضوعات پیچیده می‌باشند. آن‌ها توصیه می‌کنند که لازم است تلاش‌هایی در زمینه پیش‌بینی‌پذیری و یکدستی[۸] سرعنوان‌ها انجام پذیرد.

از سویی، پژوهش‌هایی جهت تعیین سطح هم‌خوانی – آن گونه که در زبان فارسی به انجام رسیده است– (Youngok, Ingrid, Bill, 2007) و همچنین امکان انجام نگاشت از اصطلاح‌نامه‌ها به سرعنوان‌های موضوعی کنگره (Chaplan, 1995) و از سرعنوان‌های موضوعی به سایر ابزارها از جمله فولکسونومی‌ها (Yi, Chan, 2009) انجام پذیرفته است. در لایه شبکه معنایی پژوهش‌های متنوعی پیرامون فهرست سرعنوان‌های موضوعی کنگره به انجام رسیده است (Yi, Chan, 2008; 2010; Fernandez, 1991). شباهت‌های معنایی این ابزار با برچسب‌های شبکه‌های اجتماعی مورد مطالعه قرار گرفته است(Yi, 2010) و حتی در پژوهش‌هایی اشکالات پیش آمده در روابط سلسله مراتبی در شبکه معنایی در نظام یکپارچه زبان پزشکی (یو ام ال اس) با کمک الگوریتم‌های نوشته شده و به صورت خودکار تعیین شده است (Chen, Gu, Perl, Geller, 2009a) همین پژوهشگران (Chen et al, 2009b) در

[۱] Headings with qualifier
[۲] Inverted
[۳] adjectival phrase
[۴] Prepositional Phrase
[۵] Conjunctive Phrase
[۶] Free floating phrase
[۷] Predictability
[۸] Consistent

مطالعه‌ای دیگر با استفاده از روش‌های خودکار و دستی به گروه‌بندی مفاهیم و شناسایی مفاهیمی که به اشتباه در یک حوزه وارد شده‌اند، پرداخته‌اند. کای، هینر و مگنس (Kai, Heiner, Magnus, 2007)نیز به ارزیابی ساختار شبکه‌ای اصطلاح‌نامه‌ها پرداخته‌اند و در این ارزیابی از تکنیک‌های بصری‌سازی در ترسیم درخت‌های سلسله مراتبی بهره جسته‌اند.

روش‌شناسی پژوهش

پژوهش حاضر از نوع کاربردی است و به روش تحلیل محتوا انجام پذیرفته است. در فهرست سرعنوان‌های موضوعی فارسی و کنگره، هر یک از سرعنوان‌ها، به صورت مستقل ظاهر می‌شود و پیوند میان سرعنوان‌ها به واسطه انواع روابط معنایی تعریف شده است. به این ترتیب می‌توان سرعنوان موضوعی را به منزله واحد نمونه‌گیری و ثبت تعریف کرد. سرعنوان موضوعی (فارسی و کنگره) و مجموعه روابط معنایی پیرامون آن سرعنوان به منزله واحد زمینه‌ای تعریف شد.

مقوله‌ها در رابطه باهدف‌های پژوهش تعیین و ملاک‌های هر مقوله به صورت روشن و دقیق تعریف شد. این امر، عینی و منظم بودن روش تحلیل محتوا را تضمین می‌کند و باعث تعمیم نتایج می‌شود (Sarmad, Bazargan, Hejazi, 1998) ارائه تعریف‌های دقیق و مثال‌هایی از قواعد رمزگذاری در قالب "دستور کار رمزگذاری"[1] بر پایه مبانی نظری به ویژه نظریه‌های حوزه زبان‌شناسی، تهیه شد.

در این پژوهش دو جامعه آماری حضور دارند؛ نخست فهرست سرعنوان‌های موضوعی فارسی؛ و دیگری فهرست سرعنوان‌های موضوعی کنگره. فهرست سرعنوان‌های موضوعی فارسی ویرایش سوم (Soltani, Fani, Rahbari Asl, 2002) و پیوست آن که به صورت چاپی در سال ۱۳۸۵ منتشر شده است، مبنای گردآوری داده‌ها قرار گرفت. اما از آن جایی که آخرین تغییرات سرعنوان‌ها مورد توجه پژوهش بوده است به منظور ردگیری آخرین تغییرات، داده‌های مورد نظر از طریق وب‌گاه سازمان اسناد و کتابخانه ملی ایران[2] گردآوری شد.

طبق چارچوب طراحی‌شده پژوهش و با توجه به اهداف تعیین‌شده، بخشی از سرعنوان‌های غیر موضوعی (مانند نام جغرافیایی، نام تنالگان و ...) و همچنین سرعنوان‌هایی که ساختارشان غیرطبیعی (مصنوعی) است، ضمن شمارش و ثبت تعداد[3]، از جامعه آماری کنار گذاشته شد. پس از پالایش‌های

---

۱ Coding agenda

۲ http://opac.nlai.ir/opac-prod/index.jsp

۳. پرسش دوم و سوم پژوهش ناظر بر ویژگی طبیعی یا مصنوعی بودن سرعنوان‌های موضوعی است. به این سبب، اگر چه در تحلیل ساختار زبانی سرعنوان‌ها- آن جایی که مقوله‌های دستوری و واژگانی مورد توجه است- تنها بر سرعنوان‌هایی با زبان

صورت گرفته، تعداد ۱۱۲۳۳ رکورد از مجموع ۱۸۳۹۶ سرعنوان موجود در فهرست سرعنوان‌های موضوعی فارسی ویرایش سوم و پیوست‌های آن، به منزله جامعه آماری از فهرست سرعنوان‌های موضوعی فارسی مورد مطالعه قرار گرفت. جامعه دوم پژوهش، فهرست سرعنوان‌های موضوعی کنگره[1] است که در وب‌گاه رسمی کتابخانه کنگره آمریکا دسترس‌پذیر می‌باشد. لازم به ذکر است که در این مورد تنها بر سرعنوان‌های مشترک در فهرست سرعنوان‌های موضوعی فارسی و فهرست سرعنوان‌های موضوعی کنگره تمرکز شده است. تعداد سرعنوان‌های مشترک در دو فهرست، ۹۵۱۹ رکورد تعیین شد .[2]

در پاسخگویی به چهار پرسش نخست، سرشماری انجام شد و ۱۱۲۳۳ رکورد مورد بررسی قرار گرفت. در پاسخ به پرسش‌های فرعی ۱-۱ و ۱-۲ نمونه‌گیری تصادفی از طریق نرم‌افزار آماری Spss انجام پذیرفت و به ترتیب نمونه‌ای از سرعنوان‌هایی با ساختار واژه به حجم ۳۵۷ و نمونه‌ای از سرعنوان‌های با ساختار گروه به حجم ۳۶۸ رکورد جهت بررسی انتخاب شد. پاسخگویی به پرسش پنجم، به همین شیوه، از طریق بررسی نمونه‌ای تصادفی به حجم ۳۷۶، از سرعنوان‌های مشترک در دو فهرست فارسی و کنگره، انجام پذیرفت.

ابزار و روش گردآوری داده‌ها

گردآوری داده‌ها بر اساس دستور کار کدگذاری و در قالب کتابچه‌های[3] طراحی شده در نرم‌افزار Excel انجام پذیرفت. کتابچه‌ها بر مبنای پرسش‌های پژوهش، مطالعه متون مرتبط[4] و توسط

طبیعی تمرکز شده است، اما سرعنوان‌هایی که با ساخت مصنوعی در فهرست آمده‌اند، ضمن تعیین نوع، شمارش و در محاسبه نسبت‌ها مورد استفاده قرار گرفته‌اند.

[1] http://id.loc.gov/

[2]. تعداد رکوردهای موجود در پایگاه مستند موضوع کنگره در یی و چان (Yi, Chan 2010) 341745 رکورد گزارش شده است. به منظور تعیین سرعنوان‌های مشترک میان دو فهرست موضوعی فارسی و کنگره، در فهرست پیوسته کتاب‌های نوسا، که ارتباط میان سرعنوان‌های ساخته شده در کتابخانه ملی و کتابخانه کنگره را برقرار می‌سازد، کاوشی انجام پذیرفت. به این ترتیب که هر سرعنوان در جامعه آماری، در فهرست نوسا مورد کاوش قرار گرفت در نهایت ۹۵۱۹ رکورد مشترک در این دو فهرست به دست آمد.

[3]. Sheets . این کتابچه‌ها با توجه به پرسش‌های پژوهش طراحی شد.

[4]. بخشی از این منابع عبارتند از: طباطبایی، علاءالدین (۱۳۸۸). ساختمان واژه و مقوله دستوری: تشخیص مقوله دستوری واژه‌ها، براساس ملاک‌های صرفی. تهران: پژوهشگاه فرهنگ، هنر و ارتباطات / فالک، جولیا (۱۳۷۱). زبان‌شناسی و زبان: بررسی مفاهیم بنیادی زبانشناسی. (خسرو غلامعلی‌زاده، مترجم). مشهد: آستان قدس رضوی، معاونت فرهنگی. (نشر اثر اصلی ۱۹۷۸)/

پژوهشگر طراحی شد. به منظور تأمین روایی، تناسب کتابچه‌های طراحی شده با ویژگی‌های مورد نظر در پرسش‌های پژوهش از نظر ساخت زبانی، طبق نظر متخصصان حوزه زبان‌شناسی و زبان و ادبیات فارسی مورد بازنگری قرار گرفت.

جهت تأمین پایایی در فرایند کدگذاری نیز، آن جایی که ویژگی‌های زبانی سرعنوان‌ها مورد توجه است، کدگذاری دو بار، یک بار توسط پژوهشگر و بار دیگر توسط متخصصی از حوزه زبان و ادبیات فارسی، انجام پذیرفت. همچنین برای افزایش اعتبار کار از منابعی چون لغت‌نامه دهخدا، فرهنگ معین و نرم‌افزار برچسب‌گذاری زبان فارسی[1] استفاده شد. از آن جایی که پژوهش حاضر تمام سرعنوان‌های منتشرشده در فهرست سرعنوان‌های موضوعی فارسی ویرایش سوم و پیوست آن را تحت پوشش دارد (سرشماری) و در تحلیل برخی از ویژگی‌ها از نمونه‌گیری تصادفی ساده استفاده شده است، می‌توان روایی بیرونی پژوهش را قابل‌قبول دانست.

یافته‌های پژوهش

به منظور اطلاع از آخرین تغییرات و اصلاحات انجام پذیرفته در فهرست سرعنوان‌های موضوعی، پس از شناسایی اولیه سرعنوان‌ها، هر یک از سرعنوان‌ها در فهرست پیوسته کتابخانه ملی ایران مورد کاوش قرار گرفت. همچنین از آن جایی که ارتباط این سرعنوان‌ها با سرعنوان‌های موضوعی کنگره مدنظر است، با استفاده از امکان فراهم آمده در فهرست پیوسته کتاب‌های نوسا[2] – که در جهت تحقق ایده داده‌های پیوندی در وب معنایی است و ارتباط میان سرعنوان‌های ساخته‌شده در کتابخانه ملی و کتابخانه کنگره را برقرار می‌سازد – در بازه زمانی اسفند ۱۳۸۹ تا بهمن‌ماه ۱۳۹۰، هر سرعنوان یک بار در فهرست کتابخانه ملی و بار دیگر در فهرست نوسا مورد کاوش قرار گرفت.

در جریان فرایند گردآوری داده‌ها مشخص شد که در برخی از موارد ارتباط میان سرعنوان‌های

مشکوه‌الدینی، مهدی (۱۳۷۹). توصیف و آموزش زبان فارسی. مشهد: دانشگاه فردوسی مشهد/ مشکوه‌الدینی ، مهدی (۱۳۷۳). دستور زبان فارسی: بر پایه نظریه گشتاری. (ویرایش سوم). مشهد: دانشگاه فردوسی مشهد. / وحیدیان کامیار، تقی، عمران، غلامرضا (۱۳۷۹). دستور زبان فارسی (۱). تهران: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاه‌ها (سمت). / مدرسی، فاطمه (۱۳۸۵). از واج تا جمله: فرهنگ زبان‌شناسی – دستور. تهران: چاپار./ فتوحی، محمود، عباسی، حبیب‌الله (۱۳۸۴). فارسی عمومی: درسنامه دانشگاهی. تهران: سخن.

[1]. نسخه آزمایشی تحت وب این نرم افزار در http://mohsenhosseini.com/prj/nlp/postagger/postagger_1_0.php قابل دسترس است. این نرم افزار توسط تیمی پژوهشی در دانشگاه فردوسی مشهد طراحی شده است.

[2] http://www.nosabooks.com/WebUI/Home.aspx

معادل در فهرست نوسا برقرار نشده است، علت بروز این اشکال به فعالیت ورود داده‌های فهرست سرعنوان‌های موضوعی ملی بازمی‌گردد. در ۱۱۲۳۳ رکورد مورد بررسی، ۸۱/۹ درصد آن‌ها فاقد اشکال بوده‌اند و در ۱۹/۱ درصد کاوش‌ها خطایی رخ داده است. لازم به ذکر است که در ۱۱۲۳۳ رکورد مورد بررسی ۹۲/۷ درصد از سرعنوان‌ها فاقد توضیحگر، ۵۳/۱ درصد دارای شماره رده و ۲۱/۱ درصد از آن‌ها دارای یادداشت دامنه می‌باشند.

۱. وضعیت سرعنوان‌های موضوعی فارسی از نظر تعداد و نسبت سرعنوان‌های واژگانی به سرعنوان‌های گروهی سرعنوان‌های موضوعی از نظر زبانی، به دو شکل واژه یا گروه ظاهر می‌شوند. بررسی ۱۱۲۳۳ رکورد مورد بررسی نشان‌گر آن است که ۶۳/۶ درصد از سرعنوان‌های موضوعی فارسی با ساخت گروه در فهرست ظاهرشده‌اند و ۳۶/۴ درصد از سرعنوان‌ها به صورت واژه به بازنمایی موضوع پرداخته‌اند.

۱-۱. انواع و نسبت واژه‌ها از نظر مقوله دستوری واژگانی و برخی ویژگی‌ها

به منظور تحلیل عمیق‌تر ویژگی‌های ساختاری واژه‌ها و شناسایی مقوله دستوری واژگانی از این تعداد، نمونه‌ای تصادفی به حجم ۳۵۷ انتخاب شد. در شناسایی مقوله واژگانی نمونه، از دو فرهنگ دهخدا و معین و همچنین نرم‌افزار برچسب‌زنی زبان فارسی نیز استفاده شد. آن چه قابل تأمل است، هم‌پوشانی ۴۵ درصدی میان سه ابزار زبانی مورد بررسی است. به بیانی از ۳۵۷ سرعنوان مورد بررسی، تنها ۴۵ درصد در هر سه ابزار زبانی مورد توجه، حضورداشته‌اند و این هم‌پوشانی در میان دو ابزار لغت‌نامه دهخدا و فرهنگ معین، ۶۰ درصد است.

در نمونه مورد مطالعه، بیش‌ترین سرعنوان‌ها از مقوله واژگانی اسم (۹۶/۱ ٪) بوده‌اند. درصد اندکی (۳/۴٪) صفت و ۰/۶ درصد نیز هم اسم و هم صفت گزارش‌شده‌اند. آن گونه که انتظار می‌رود هیچ سرعنوانی در مقوله قید، حرف و فعل شناسایی نشد. همین واژه‌ها از نظر ساختمان واژه مورد بررسی قرار گرفتند. مشخص گردید که در نمونه مورد بررسی ۶۲/۲ درصد از سرعنوان‌های با ساختار واژه، ساختمانی ساده دارند. ۸/۱ درصد از نظر ساختمان واژه، مرکب، ۹ درصد مشتق و ۲۰/۷ درصد از آن‌ها مشتق – مرکب می‌باشند.

از مجموعه نمونه مورد بررسی در این بخش، تعداد قابل‌ملاحظه‌ای از سرعنوان‌ها (۳۴۳ از ۳۵۷ سرعنوان) در مقوله اسم قرار دارند. این سرعنوان‌ها از نظر جمع و مفرد بودن مورد بررسی قرار گرفتند. نسبت بیشتری از سرعنوان‌هایی که در مقوله اسم هستند، مفرد و تنها ۳۰/۳ درصد از سرعنوان‌ها، با نشانه

جمع و یا به صورت جمع مکسر ظاهرشده‌اند.

۲-۱ .تنوع و نسبت انواع واحدهای زبانی گروه در فهرست سرعنوان‌های موضوعی فارسی

گفته شد که بخش قابل‌توجهی (۶۳/۶ ٪) از سرعنوان‌های موضوعی فارسی با ساختار گروهی ساخته‌شده‌اند. همانند تحلیلی که پیرامون واژه‌ها انجام پذیرفت، نمونه‌ای تصادفی از سرعنوان‌های گروهی به حجم ۳۶۸ انتخاب شد و نوع گروه هر یک از سرعنوان‌ها تعیین گردید. جدول شماره ۱ فراوانی و درصد انواع گروه‌های زبانی را نشان می‌دهد.

جدول شماره ۱. مقوله دستوری گروهی سرعنوان‌های موضوعی (n=۳۶۸)

مقوله دستوری گروهی		فراوانی	درصد
گروه اسمی	اضافی	۵۸	۸/۱۵
	وصفی	۲۲۰	۸/۵۹
	عطفی	۱۵	۱/۴
	مقید	۲۷	۳/۷
گروه صفتی		۰	۰/۰
گروه قیدی		۰	۰/۰
گروه حرف اضافه‌دار		۰	۰/۰
گروه فعلی		۰	۰/۰
ترکیبی از چند گروه		۴۸	۰/۱۳
مجموع		۳۶۸	۱۰۰

آن گونه که در جدول مشخص شده است، سرعنوان‌های موضوعی نمونه، فاقد ساخت گروه‌های صفتی، قیدی، حرف اضافه‌دار و فعلی هستند. گروه اسمی بیش‌ترین ساخت گروهی (۸۷ درصد) را به خود اختصاص داده است و ۱۳ درصد سرعنوان‌های نمونه مورد بررسی، ترکیبی از چند گروه (دو و بیشتر) هستند. نکته قابل‌توجه آن است که این ترکیب‌ها نیز از چند سر عنوان گروه اسمی ساخته‌شده‌اند. این در حالی است که در فهرست سرعنوان‌های موضوعی کنگره وجود انواع سرعنوان‌های گروه صفتی، و گروه حرف اضافه‌دار نیز گزارش شده است .(Yi, Chan, 2010)

۲. نسبت ساختار زبان طبیعی به زبان مصنوعی

به طور کلی می‌توان سرعنوان‌های موضوعی را در دو دسته زبان طبیعی و زبان مصنوعی دسته‌بندی کرد. در نظام موضوعی سرعنوان‌های موضوعی فارسی، آن گونه که در سرعنوان‌های مورد بررسی (۱۱۲۳۳ رکورد) مشاهده می‌شود، ۸۶/۱۷ درصد سرعنوان‌ها، زبان طبیعی محسوب می‌شوند و تنها ۱۳/۸۳

درصد، مصنوعی‌اند. تعداد قابل‌توجهی از سرعنوان‌های مصنوعی (۱۲/۱۶٪) به واسطه داشتن تقسیم فرعی، در این گروه شمارش‌شده‌اند. در فهرست سرعنوان‌های موضوعی کنگره حدود ۶۷/۷ درصد از سرعنوان‌ها بدون تقسیم فرعی هستند و ۳۲/۳ درصد از سرعنوان‌ها به همراه تقسیمات فرعی ظاهرشده‌اند (Yi, Chan, 2010).

۳. تعیین سطح پیش‌همارایی و نسبت پیش‌همارایی طبیعی به پیش‌همارایی مصنوعی در سرعنوان‌های موضوعی فارسی

در بررسی ۱۱۲۳۳ رکورد سرعنوانی مشاهده می‌شود که ۳۶/۴ درصد از سرعنوان‌ها، تک‌کلمه‌ای و ۶۳/۶ درصد از آن‌ها ترکیبی از دو کلمه و بیشترند. سرعنوان‌های دو کلمه‌ای با ۴۲/۱ درصد شکل غالب سرعنوان‌های موضوعی در سرعنوان‌های مورد مطالعه است. هر چه بر تعداد واژه‌ها و به بیانی طول سرعنوان افزوده می‌شود، تعداد سرعنوان‌ها کاهش می‌یابد. طولانی‌ترین سرعنوان‌های موضوعی فارسی، ۶ واژه‌ای هستند و حدود ۰/۲ درصد از سرعنوان‌ها را به خود اختصاص داده‌اند.

طبق تعریف، سطح پیش‌همارایی سرعنوان‌ها یا تعداد متوسط واژه‌ها در هر سرعنوان موضوعی با محاسباتی مشابه تعیین میانگین تعداد واژه‌ها در کل سرعنوان‌ها انجام می‌پذیرد. در فهرست سرعنوان‌های موضوعی کنگره وضعیت تقریباً مشابهی برقرار است به گونه‌ای که سرعنوان‌های تک کلمه‌ای ۱۲/۶ درصد و سرعنوان‌های دو کلمه‌ای بیش‌ترین رخداد یعنی ۳۶/۹ درصد را به خود اختصاص داده‌اند. در این مورد، همچون فهرست سرعنوان‌های موضوعی فارسی، با افزایش تعداد کلمات، رخداد سرعنوان‌ها سیر نزولی می‌یابد (Yi, Chan, 2010).

در این سطح از پیش‌همارایی در سرعنوان‌های موضوعی مورد بررسی، آن چه قابل توجه است، تعداد ۷۹/۸۷ درصدی پیش‌همارایی طبیعی در مقابل ۲۰/۱۴ درصد پیش‌همارایی مصنوعی[1] و به این ترتیب نسبت ۳/۹۶ برابری پیش‌همارایی طبیعی به مصنوعی است.

۴. ویژگی‌های شبکه معنایی حاکم بر سرعنوان‌های موضوعی فارسی به لحاظ پیوند میان سرعنوان‌ها

---

[1]. به منظور تعیین پیش‌همارایی مصنوعی سرعنوان‌هایی که دارای تقسیم فرعی هستند، سرعنوان‌های مقلوب، و آن دسته از سایر اشکال سرعنوان‌هایی که در گروه زبان مصنوعی قرار گرفته و بیش از یک واژه دارند، در نظر گرفته شده و شمارش شده‌اند.

۱-۴. ساختار شبکه جزئی روابط معنایی سرعنوان‌های موضوعی فارسی

در ۱۱۲۳۳ رکورد سرعنوانی مورد بررسی، تعداد (۸۹/۹۹ ٪) ۱۰۱۰۹ شبکه جزئی روابط معنایی[۱] (LRS) شناسایی شد و تعداد (۱۰/۰۱٪) ۱۱۲۴ مورد از سرعنوان‌ها، فاقد هر گونه پیوندی با سایر سرعنوان‌ها می‌باشند.

همانگونه که در جدول شماره ۱ آمده است. بیش‌ترین نوع رابطه در فهرست سرعنوان‌های موضوعی فارسی، رابطه هم‌ارزی با ۸۲ درصد و میانگین ۲/۲۶ و کمترین نوع رابطه، رابطه هم‌بسته با ۲ درصد و میانگین ۰/۰۵ می‌باشد.

شمار مفهومی دو بعدی است، به این معنی که با وجود و یا عدم وجود سروکار دارد. مقدار شمار، تعداد واقعی کاربرد هر رابطه را در ۱۱۲۳۳ رکورد نشان می‌دهد. به این ترتیب مشخص است که به طور مثال در ۹۷۱۰ رکورد از ۱۱۲۳۳ رکورد (۸۶/۴۴٪) رابطه هم‌ارزی وجود داشته است. درصد شمار، درصد واقعی کاربرد یک رابطه در شبکه جزئی روابط معنایی سرعنوان‌های مورد بررسی است. مشاهده می‌شود که ۸۹/۹۹ درصد از سرعنوان‌ها حداقل از یک نوع رابطه معنایی برخوردارند و کمترین نوع رابطه معنایی مورد استفاده در سرعنوان‌ها، رابطه هم‌بسته است. در مجموع تعداد ۳۲۳۷۴ رابطه معنایی میان ۱۱۲۳۳ رکورد مورد مطالعه برقرار است و به طور متوسط هر سرعنوان، دارای ۲/۸۸ رابطه با سایر سرعنوان‌هاست.

جدول ۱. خلاصه معیارهای اندازه‌گیری روابط مطرح در سرعنوان‌های موضوعی فارسی مورد مطالعه

معیارها	تعداد	درصد	اندازه‌های گرایش مرکزی		اندازه‌های پراکندگی			شمار	درصد شمار[۲]
نوع رابطه			میانگین	میانه	حداقل	حداکثر	انحراف معیار		
هم‌ارزی	۲۶۰۴۷۰	۸۲	۳۶/۲	۲	۰	۱۸	۰۸۰/۲	۹۷۱۰	۴۴/۸۶
اخص	۳۷۰۳	۱۱	۳۳/۰	۰۰/۰	۰	۴۹	۰۹۴/۱	۲۰۴۷	۴۶/۱۸
اعم	۱۵۸۴	۵	۱۴/۰	۰۰/۰	۰	۶	۴۲۲/۰	۱۳۳۴	۸۷/۱۱

---

[۱]. هر سه نوع رابطه معنایی هم‌ارزی، سلسله مراتبی و هم‌بسته که به ترتیب در فهرست سرعنوان‌های موضوعی فارسی با نشانه‌های اختصاری X، XX، ن.ن و م. آمده اند، را در بر می‌گیرد. در این شبکه، تمام سرعنوان‌هایی که به صورت مستقیم ذیل یک سرعنوان آمده‌اند –صرف‌نظر از این که پذیرفته شده یا پذیرفته نشده باشند – مورد توجه‌اند.

[۲]. درصد شمار به نسبت ۱۱۲۳۳ رکورد محاسبه شده است. به طور نمونه ۹۷۱۰ سرعنوان از ۱۱۲۳۳ سرعنوان موضوعی دارای رابطه هم‌ارز بوده است که درصد آن ۸۶/۴۴ درصد خواهد شد.

هم‌بسته	۶۱۶	۲	۰۵/۰	۰۰/۰	۰	۸	۲۵۷/۰	۵۶۴	۰۲/۵
تمام روابط	۳۲۳۷۴	۱۰۰	۸۲/۲[1]	۲	۰	۵۸	۵۷۵/۲	۱۰۱۰۹	89/99

۲-۴. چگونگی نسبت پیوستگی و دو مکمل آن یعنی نسبت وابستگی و انسجام[2] میان سرعنوان‌ها در فهرست سرعنوان‌های موضوعی فارسی

در ارزیابی شبکه معنایی اصطلاح‌نامه‌ها، مقیاس‌هایی کمی مورد استفاده قرار می‌گیرد. این نسبت‌ها، به اعتبار قبول ضمنی انواع روابط معنایی در فهرست سرعنوان‌های موضوعی فارسی، در مورد این ابزار بازنمون دانش محاسبه شد.

جدول شماره ۲ . نسبت پیوستگی و دو مکمل آن یعنی نسبت وابستگی و انسجام

معیارهای کمی ارزیابی در فهرست سرعنوان‌های موضوعی فارسی		
نسبت پیوستگی	نسبت انسجام	نسبت وابستگی
۲۹۶/۰	۲۷۰/۰	۰۵۰/۰

آن‌گونه که در جدول شماره ۲ نمایش داده شده است بیشترین نسبت محاسبه شده، همان‌گونه که انتظار می‌رود، به نسبت پیوستگی تعلق یافته است. در محاسبه این نسبت، سرعنوان‌هایی که دارای یکی از انواع روابط سلسله مراتبی و هم‌بسته باشند شمارش شده و بر تعداد سرعنوان‌ها (در این جا **۱۱۲۳۳** رکورد) تقسیم می‌شود، منطقی به نظر می‌رسد که عدد حاصل شده برای این نسبت بیش از سایر نسبت‌ها باشد. پس از آن نسبت انسجام که تحت تأثیر سرعنوان‌های دارای رابطه سلسله مراتبی است، قرار گرفته است و نسبت وابستگی، که تنها تحت تأثیر رابطه معنایی هم‌بسته قرار دارد، کمترین میزان را به خود اختصاص داده است.

---

۱. عدد مربوط به اندازه‌های گرایش مرکزی و پراکندگی، شمار و درصد آن در ردیف کل روابط معنایی، به طور جداگانه محاسبه شده است و مجموع عددهای ردیف‌های بالای آن ستون نیست.

۱. نسبت پیوستگی: نسبت پیوستگی یا ارتباط عبارت است از نسبت سرعنوان‌های دارای ارجاع‌های متقابل (دارای ارتباط سلسله مراتبی و هم‌بسته) به تعداد کل سرعنوان‌ها. نسبت وابستگی: عبارت است از نسبت سرعنوان‌هایی که دارای سرعنوان هم‌بسته‌اند به تعداد کل سرعنوان‌ها. این نسبت، مکمل نسبت پیوستگی است. نسبت انسجام: عبارت است از نسبت سرعنوان‌هایی که دارای رابطه اعم و اخص (سلسله مراتبی) هستند به تعداد کل سرعنوان‌ها. این نسبت، مکمل نسبت پیوستگی است (Amirhosseini, 2007a)

۵. فهرست سرعنوان‌های موضوعی فارسی در مقایسه با کنگره

در **۱۱۲۳۳** رکورد مورد بررسی، تعداد قابل ملاحظه‌ای از سرعنوان‌ها – **۸۴/۷** درصد– در هر دو فهرست موضوعی فارسی و کنگره مشترک است. اخوتی و دیگران (Okhovati et al, 2009) این انطباق را در حوزه کتابداری و اطلاع‌رسانی میان دو فهرست سرعنوان‌های موضوعی فارسی و کنگره، **۷۱/۶۷** درصد و بیش از سرعنوان‌های موضوعی سیرز (۳۲/۰۸٪) و لیزا (۵۹/۵۸٪) به دست آورده است. روشن ساختن شباهت‌ها و تفاوت‌های میان این دو ابزار به واسطه نمونه‌ای **۳۷۶** تایی از **۹۵۱۹** رکورد مشترک در دو فهرست موضوعی انجام پذیرفت.

نتایج آزمون آماری ویلکاکسون[1] نشان‌گر وجود تفاوت معنادار میان دو ابزار سرعنوانی در سه ویژگی ۱. نوع زبان (طبیعی یا مصنوعی) (P-value= 0.000) ، ۲. به کارگیری یادداشت دامنه (P-value= 0.003) و ۳. ارتباط با رده‌بندی کنگره (P-value= 0.000) است. بهره‌گیری از ساختار زبان طبیعی و افزودن یادداشت دامنه در فهرست فارسی بیش از فهرست کنگره است. این در حالی است که به واسطه ارائه شماره رده، ارتباط میان سرعنوان کنگره و رده‌بندی کنگره، بیش از فهرست فارسی و رده‌بندی کنگره برقرار شده است.

۵-۱. مقایسه ساختار شبکه جزئی روابط معنایی در دو فهرست سرعنوان‌های موضوعی فارسی و کنگره

جدول شماره ۳ تفاوت رویکرد در ساخت روابط معنایی دو ابزار بازنمون دانش را منعکس می‌سازد.

جدول ۳. اطلاعات توصیفی روابط معنایی حاکم بر سرعنوان های موضوعی فارسی و کنگره (**n= ۳۷۶**)

معیارهای اندازه‌گیری	تعداد		درصد		شمار		درصد شمار[2]	
نوع رابطه	فارسی	کنگره	فارسی	کنگره	فارسی	کنگره	فارسی	کنگره
هم‌ارزی	۹۴۹	۶۷۳	۷/۸۲	۱/۱۹	۳۲۱	۲۷۹	۳۷/۸۵	۲۰/۴۷
اخص	۱۲۵	۲۱۲۹	۹/۱۰	۶/۶۰	۸۱	۲۰۸	۵۴/۲۱	۳۲/۵۵
اعم	۴۸	۵۴۸	۲/۴	۶/۱۵	۴۸	۳۲۱	۷۶/۱۲	۳۷/۸۵
هم‌بسته	۲۵	۱۶۵	۲/۲	۴/۷	۲۶	۱۱۲	۹۱/۶	۷۸/۲۹
تمام روابط معنایی	۱۱۴۷	۳۵۱۵	۱۰۰	۱۰۰	۳۴۳	۳۶۶	۲۲/۹۱	۳۴/۹۷

[1] Wilcoxon

[2] با احتساب تعداد ۳۷۶ سرعنوان محاسبه شده است.

چنین به نظر می‌رسد که بیشترین تلاش کتابخانه ملی ایران در ساخت رابطه هم‌ارز بوده است به طوری که ۸۲/۷ درصد از کل روابط معنایی تعریف شده در این فهرست از نوع رابطه هم‌ارز است. این در حالی است که رابطه اخص با ۶۰/۶ درصد، بیشترین نوع رابطه در سرعنوان‌های موضوعی کنگره می‌باشد. در ستون شمار، که کاربرد واقعی رابطه را نشان می‌دهد، عدم توازن در توجه به انواع پیوندهای معنایی در نمونه مورد بررسی، قابل مشاهده است. با اندک دقتی در این ستون مشخص می‌شود که رابطه هم ارز در ۳۲۱ مورد از ۳۷۶ رکورد مورد بررسی (۸۵/۳۷٪) در فهرست سرعنوان‌های موضوعی فارسی مورد استفاده قرار گرفته است و اگر چه درصد سرعنوان‌هایی با رابطه اخص بیش از سایر انواع روابط در فهرست کنگره بوده است، اما شمار روابط در این فهرست آشکار می‌سازد که کاربرد رابطه اعم در سطح سرعنوان‌ها، بیش از سایر روابط است و در ۳۲۱ مورد از رکوردهای مورد بررسی رخ داده است. به بیانی دیگر رخداد رابطه اعم بیش از سایر روابط در نمونه مورد بررسی از سرعنوان‌های کنگره است. میانگین روابط معنایی به ترتیب در دو فهرست فارسی و کنگره در مورد رابطه هم‌ارزی ۲/۵۲ و ۱/۷۹، در رابطه اخص ۰/۳۳ و ۵/۶۶، در رابطه اعم ۰/۱۳ و ۱/۴۶ و در رابطه هم‌بسته ۰/۰۷ و ۰/۴۴ است. در کل میانگین تعداد کل روابط در فهرست فارسی ۳/۰۵ و در کنگره ۹/۳۵ می‌باشد. بالاترین تعداد رابطه در فهرست کنگره به رابطه اخص با ۱۳۷ مورد و بالاترین تعداد رابطه در فهرست فارسی به رابطه هم‌ارزی با ۱۴ مورد اختصاص یافته است. این یافته نشانگر توجه به اخص‌گرایی و ساخت رابطه اخص در فهرست سرعنوان‌های موضوعی کنگره می‌باشد. نتیجه آزمون آماری T دوسویه[1] تفاوتی معنادار و قوی (P-value=0.000) از نظر تعداد روابط معنایی، در مورد تک تک روابط در دو ابزار مورد مطالعه نشان می‌دهد.

۲-۵. مقایسه نسبت پیوستگی و دو مکمل آن یعنی نسبت وابستگی و انسجام سرعنوان‌های موضوعی در دو فهرست

پیش‌تر در جدول ۲ مقیاس‌های کمی نسبت پیوستگی، نسبت وابستگی و انسجام برای ۱۱۲۳۳ رکورد سرعنوانی مورد مطالعه محاسبه و گزارش شد. در جدول ۴ این نسبت‌ها به طور مجزا در نمونه ۳۷۶ تایی سرعنوان‌های مشترک انتخاب شده از دو فهرست سرعنوان موضوعی کنگره و فارسی آمده است. ملاحظه می‌شود که نسبت‌های محاسبه شده در سرعنوان‌های فارسی نمونه، با تفاوتی قابل چشم‌پوشی – به جز نسبت پیوستگی – مشابه نسبت‌های به دست آمده در مورد ۱۱۲۳۳ رکورد سرعنوان موضوعی فارسی (جدول ۲) به دست آمده است. اما در مقایسه دو ابزار بازنمون، وضعیت سرعنوان‌های موضوعی کنگره در

[1] Paired T test

نمونه ۳۷۶ تایی، تفاوت‌های قابل ملاحظه‌ای را با سرعنوان‌های معادل آن در فارسی نشان می‌دهد. این تفاوت‌ها ناشی از تفاوت رویکرد مستندسازان کتابخانه ملی آمریکا و کتابخانه ملی ایران در گسترش روابط معنایی – به ویژه روابط سلسله مراتبی و هم‌بسته – است و پیش‌تر در تفاوت درصد کاربرد روابط معنایی (جدول ۳)، تفاوت میانگین‌ها و به طور منطقی تفاوت معنادار در استفاده از روابط معنایی در دو ابزار موضوعی جلوه‌گر شده است.

جدول ۴. مقایسه نسبت پیوستگی و دو مکمل آن یعنی نسبت وابستگی و انسجام در دو فهرست (n= ۳۷۶)

شاخصه‌ها	سرعنوان موضوعی فارسی	سرعنوان موضوعی کنگره	نتایج آزمون نسبت	
			آماره Z	مقدار P-value
نسبت پیوستگی	۳۵۳/۰	۸۹۸/۰	۹۲/۱۵	۰۰۰/۰
نسبت انسجام	۳۰۰/۰	۸۸۸/۰	۴۲/۱۶	۰۰۰/۰
نسبت وابستگی	۰۶۳/۰	۲۹۲/۰	۱۹/۸	۰۰۰/۰

نسبت پیوستگی سرعنوان‌های موضوعی کنگره ۲/۶۸ برابر نسبت پیوستگی سرعنوان‌های موضوعی فارسی است (۰/۸۹۸ در مقابل ۰/۳۳۵). نسبت انسجام سرعنوان‌های موضوعی کنگره نیز بزرگتر از سرعنوان‌های موضوعی فارسی و ۲/۹۶ برابر سرعنوان‌های فارسی معادل آن است (۰/۸۸۸ در مقابل ۰/۳۰۰). این واقعیت، تثبیت‌کننده کاربرد بیشتر روابط سلسله مراتبی و هم‌بسته در فهرست کنگره به نسبت فهرست سرعنوان‌های فارسی است. همان‌طور که در جدول ۴ به وضوح مشخص است، با حذف سرعنوان‌های دارای رابطه هم‌بسته در محاسبه نسبت انسجام، این نسبت در هر دو فهرست فارسی و کنگره تفاوت اندکی در مقایسه با نسبت پیوستگی در همان فهرست‌ها یافته است (در فارسی از ۰/۳۳۵ به ۰/۳۰۰ و در کنگره از ۰/۸۹۸ به ۰/۸۸۸ کاهش یافته است). نسبت وابستگی، که تحت تاثیر رابطه هم‌بسته است، در فهرست کنگره ۴/۶۳ برابر فهرست فارسی است. یافته‌ها نشانگر آن است که تعریف روابط معنایی در فهرست سرعنوان‌های کنگره در مقایسه با فهرست فارسی از غنای بیشتری برخوردار است و تفاوتی معنادار میان سرعنوان‌های موضوعی فارسی و کنگره از نظر نسبت‌های پیوستگی، انسجام و وابستگی برقرار است.

سخن پایانی

در مقدمه فهرست سرعنوان‌های موضوعی فارسی دستورالعمل‌هایی چون بهره‌گیری از زبان طبیعی (Soltani, Fani, Rahbari Asl, 1994) از جمله عدم استفاده از فرم مقلوب و به کارگیری فرم تقسیم فرعی (p28) و در عین حال ارجحیت شکل عبارتی سرعنوان نسبت به تقسیم فرعی[1](p28)، افزودن یادداشت و دستورالعمل به منظور کاستن از هر گونه ابهام (p19, p26)و استفاده به جا از توضیحگر در اصطلاحاتی که مشترک لفظی هستند (p14) مورد اشاره قرار گرفته است. همچنین از نظر ساختمان، سرعنوان‌های ساده و تک‌واژه‌ای را بهترین نوع سرعنوان دانسته است (p28). از نظر مقوله دستوری – واژگانی و گروهی – دستورالعملی در مقدمه نیامده اما با توجه به مثال‌های ذکر شده و انواع سرعنوان‌هایی که به لحاظ شکلی در فهرست آمده است، به طور ضمنی مشخص می‌شود که مقوله واژگانی اسم و مقوله گروهی گروه‌های اسمی مورد نظر بوده‌اند (p28). هیچ یک از دو شکل جمع یا مفرد سرعنوانی در فهرست مورد تأکید نیست و تنها قواعد مختصری پیرامون برخی از موضوعات که لازم است جمع یا مفرد بیایند، ارائه شده است (p29, p30). به این ترتیب نمی‌توان ارجحیتی از این باب قائل شد و تنها بر اهمیت یکدستی در انتخاب سرعنوان‌های جمع یا مفرد و به کارگیری آن در کل فهرست تأکید شده است.

تحلیل یافته‌ها در ۱۱۲۳۳ رکورد مورد بررسی، نشانگر آن است که ۹۲/۷ درصد از سرعنوان‌ها فاقد توضیحگر، ۵۳/۱ درصد دارای شماره رده و ۲۱/۱ درصد از آن‌ها دارای یادداشت دامنه می‌باشند. بیشتر سرعنوان‌های ساخته شده در فهرست سرعنوان‌های موضوعی فارسی از نظر مقوله دستوری در مقوله گروه می‌گنجند و شکل سرعنوان‌های دو واژه‌ای (۴۲/۱٪) و گروه اسمی (۸۷٪) غالب است. در سرعنوان‌های تک‌کلمه‌ای یا تک‌واژه‌ای، غالب سرعنوان‌ها، اسم (۹۶/۱٪)، ساده (۶۲/۲٪) و مفرد (۶۹/۷٪) است. این سرعنوان‌ها بیشتر با ساختار زبان طبیعی (۸۶/۱۷٪) هماهنگ‌اند و از نظر پیش‌همارایی نیز درصد پیش‌همارایی طبیعی به نسبت پیش‌همارایی مصنوعی بیشتر است. ویژگی‌های حاصل شده از تحلیل‌ها، نشان‌گر پایبندی نسبی و مطلوب سرعنوان‌ها از دستورالعمل‌های تعریف شده در مقدمه فهرست سرعنوان‌های موضوعی فارسی است. در بررسی تفاوت میان دو فهرست فارسی و کنگره، تفاوتی معنادار میان دو فهرست فارسی و کنگره از نظر ۱. طبیعی یا مصنوعی بودن زبان، ۲. به کارگیری یادداشت دامنه و ۳. ارتباط با رده‌بندی کنگره مشاهده گردید. همچنین مشخص شد که فهرست سرعنوان‌های موضوعی فارسی از نظر زبان، بیش از کنگره پایبند به اصل طبیعی بودن سرعنوان‌ها بوده است.

---

[1].اما در ادامه اشاره می‌کند که نمی‌توان به طور قطع گفت که فرم عبارتی همیشه بر فرم تک‌واژه‌ای با تقسیم فرعی ارجحیت دارد یا بالعکس (۲۸p).

در باب روابط معنایی، تعداد روابط معنایی تعریف شده در فهرست سرعنوان‌های فارسی در مقایسه با کنگره بسیار اندک است و تفاوتی معنادار میان این دو از نظر گسترش شبکه معنایی حاکم است. بیشترین روابط معنایی تعریف شده در فهرست سرعنوان‌های فارسی، رابطه هم‌ارز و در کنگره، رابطه اخص است. به این ترتیب چنین به نظر می‌رسد که شبکه معنایی سرعنوان‌های کنگره ضمن غنی بودن، به نسبت سرعنوان‌های فارسی، اخص‌تر نیز می‌باشد. محاسبه نسبت‌های دسترسی، پیوستگی و دو مکمل آن یعنی وابستگی و انسجام نیز تفاوتی میان دو فهرست را نشان می‌دهد. از آن جایی که شبکه روابط معنایی در کنگره از نظر تعداد، غنی‌تر از فهرست فارسی است، نسبت‌های به دست آمده برای این دو ابزار فاصله زیادی با یکدیگر دارند.

اخوتی و دیگران (Okhovati et al, 2009) با مقایسه ساختار شبکه مفهومی[1] سرعنوان‌های موضوعی فارسی در حوزه کتابداری و اطلاع‌رسانی با شبکه مفهومی سه ابزار بین‌المللی فهرست سرعنوان‌های موضوعی کنگره، فهرست سرعنوان‌های موضوعی سیرز و چکیده‌نامه کتابداری و اطلاع‌رسانی (لیزا) نشان داد که ساختار شبکه مفهومی در سرعنوان‌های موضوعی فارسی از نظر روابط واژگانی، انسجام و پیوستگی[2] لازم را ندارد و شبکه مفهومی کتابداری و اطلاع‌رسانی حاکم بر سرعنوان موضوعی فارسی با شبکه مفهومی مرتبط با کتابداری و اطلاع‌رسانی در سرعنوان‌های کنگره و سیرز و همچنین لیزا تفاوت معناداری داشته است. وی خاطر نشان ساخته است که از نظر شکلی نیز در شبکه مفهومی این سرعنوان، نوعی از بهم‌ریختگی مشاهده می‌شود و بر این پایه نتیجه گرفته است که به احتمال بتوان گفت پایه نظری سرعنوان موضوعی فارسی چندان مستحکم نیست و ترسیم روابط واژگانی در آن تابع الگویی منطقی و عملی نیست. از سویی خاجویی (Khojoui, 1993) در بررسی دو اصطلاح‌نامه متالوژی ASM و SDIM با فرایندی شبیه به تشکیل شبکه جزئی روابط و محاسبه ارزش هر یک از توصیفگرهای اصطلاح‌نامه بر مبنای تعداد روابط اعم و اخص آن به این نتیجه رسیده است که هر چه روابط متقابل بین اصطلاحات یک اصطلاح‌نامه بیشتر باشد، اصطلاحات آن در موقعیت مستحکم‌تری

[1]. منظور مجموعه‌ای از گره‌های معنایی و ارتباط میان آن‌هاست. در پژوهش حاضر این شبکه مفهومی با عنوان شبکه معنایی و اختصاصا شبکه جزئی روابط، مورد بررسی قرار گرفته است. تاکید می‌شود که نظام مفهومی مورد بررسی در این پژوهش با آن چه اخوتی و دیگران (Okhovati et al, 2009) از شبکه مفهومی در نظر گرفته است، تفاوت دارد.

[2]. انسجام و پیوستگی مورد نظر اخوتی و دیگران (Okhovati et al, 2009) قضاوتی بر مبنای شبکه‌های ترسیم شده است و محاسباتی برای آن صورت نپذیرفته است لذا متفاوت از سنجه‌های کمی است که در پژوهش حاضر و بر اساس فرمول‌های مشخص محاسبه شده است.

خواهند بود. با توجه به این دو مطلب، چنین به نظر می‌رسد که سرعنوان‌های موضوعی فارسی در این نظام موضوعی، از موقعیت چندان مستحکمی برخوردار نیستند.

با توجه به یافته‌های پژوهش مشخص می‌شود که اگر چه سرعنوان‌های موضوعی فارسی از نظر ساختار زبانی، وضعیت به نسبت مطلوبی دارد و سعی در رعایت ضوابط مندرج در مقدمه فهرست سرعنوان‌ها از جمله ساخت طبیعی داشته است، اما عدم توجه کافی به گسترش روابط معنایی، وجود اشکالات در تعریف این شبکه و حتی اشکالات حاصل شده در مرحله ورود اطلاعات به پایگاه کتابخانه ملی ایران لزوم بازنگری در فهرست را پررنگ می‌کند. لزوم این نوع از بازنگری تنها با این پیش‌فرض توجیه‌پذیر است که ساختار فعلی سرعنوان‌ها و روند حاکم بر توسعه آن از سوی جامعه علمی و حرفه‌ای ایران پذیرفته شده باشد و تصمیمات مبنایی‌تری چون حرکت به سمت ساختار اصطلاحنامه‌ای و گسترش مجزای سرعنوان‌های موضوعی در حوزه‌های موضوعی گوناگون با رویکردهایی چون رویکرد تحلیل حوزه، آن گونه که در اخوتی و دیگران (Okhovati et al, 2009) پیشنهاد شده است، مطرح نباشد.

کتابنامه

اخوتی، مریم؛ داورپناه، محمدرضا؛ دیانی، محمدحسین؛ شعبانی ورکی، بختیار (۱۳۸۸). تحلیل و مقایسه شبکه مفهومی حاکم بر سرعنوان موضوعی و متون فارسی کتابداری و اطلاع رسانی بر اساس رویکرد تحلیل حوزه. پژوهش‌نامه مبانی تعلیم و تربیت, ۱۰. (۲) ، ص ۵-۲۶. بازیابی شده در ۱۳۸۹ /۴/۲ از http://jm.um.ac.ir/index.php/psychology/article/view/1594

آخوندی، شکوفه، حری، عباس (۱۳۸۵). تحلیل ساختاری توصیفگرهای فارسی پایگاه های اطلاعاتی پژوهشگاه اطلاعات و مدارک علمی ایران. کتابداری، ۴۰(۴۵)، ص ۳۷-۶۲.

امیرحسینی، مازیار (۱۳۸۶ الف). ارزیابی عوامل کمی و کیفی در اصطلاحنامه‌های فارسی. کتاب ماه کلیات، ۱۱۸، ص ۱۰۰-۱۰۴.

امیرحسینی، مازیار (۱۳۸۶، آبان ۲۴). ارزیابی کمی مقوله پیوستگی در شبکه توصیفگرهای اصطلاحنامه اصول فقه، منطق، فلسفه اسلامی، علوم قرانی و کلام. مقاله ارائه شده در همایش ملی اصطلاحنامه ها و کاربردهای آن در محیط الکترونیکی، قم.

امیرحسینی، مازیار (۱۳۸۴). ارزیابی کمی مفاهیم تعداد و شمار روابط اصطلاحنامه‌ای در شبکه اصطلاحات اصطلاحنامه کشاورزی کب. فصلنامه کتاب، ۱۶(۱۴)، ص ۱۳-۳۲.

امیرحسینی، مازیار (۱۳۸۶ ب). ارزیابی کمی نسبت تعادل در اصطلاحنامه های علوم اسلامی حرکت به سمت تحلیل توانمندی واژگان مدخل. کتاب ماه کلیات، ۱۱۸، ص ۱۶-۲۹.

امیرحسینی، مازیار (۱۳۸۶ ج). ارزیابی کیفی اصطلاحنامه های اصفا و شیمی بر مبنای دیدگاه های محتوایی، ساختاری، فیزیکی و اداره اصطلاحنامه. کتاب ماه کلیات، ۱۱۸، ص ۶۵-۷۲.

بزرگی، اشرف السادات (۱۳۷۰). بررسی میزان سازگاری موضوع ها و عنوان ها در کتاب های فارسی. فصلنامه کتاب، ۲(۱-۴)، ص ۱۱۴-۱۴۵.

خاجویی، فائزه (۱۳۷۲). بررسی میزان قوت و ضعف روابط سلسله مراتبی در اصطلاحنامه با نگرشی بر دو اصطلاحنامه متالوژی ASM و SDIM. پیام کتابخانه، ۳(۲و۳)، ص ۲۲-۳۳.

خانی، صدیقه؛ علی بیک، محمدرضا و حقانی، حمید (۱۳۸۶) مقایسه ساختار اصطلاحنامه های پایگاه‌های اطلاعاتی Pubmed و Embase با استاندارد اصطلاحنامه نویسی سازمان ملی استانداردهای اطلاعاتی آمریکا و بررسی شیوه های نمایه سازی دو پایگاه. مدیریت سلامت، ۱۰(۲۷)، ص ۲۷- ۳۲.

زهادی، فیروزان (۱۳۸۸) بررسی میزان قانون مندی و یکدستی شکلی و محتوایی سرعنوان های همراه با تقسیم فرعی یا توضیح‌گر اسلام، یا ترکیب [موضوع] و اسلام در بانک مستند سرعنوانهای موضوعی کتابخانه ملی، در مقایسه با سرعنوان های مشابه در بانک مستند موضوعی کتابخانه کنگره آمریکا. کتابداری و اطلاع رسانی . (12) 4 . بازیابی‌شده‌در1/4/91،از http://www.aqlibrary.org/index.php?module=TWArticles&file=index&func=view_pubarticles&did=743&pid=۱۰

سرمد، زهره، بازرگان، عباس، حجازی، الهه (۱۳۷۷). روش‌های تحلیق در علوم رفتاری. تهران: آگاه.

سلطانی، پوری؛ فانی، کامران؛ رهبری اصل، مهناز (۱۳۸۱). سرعنوان‌های موضوعی فارسی (ویرایش۳). تهران: سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران.

سلطانی، پوری؛ فانی، کامران؛ زهادی، فیروزان (۱۳۸۲). سرعنوان‌های موضوعی فارسی ویراست سوم - پیوست یک. تهران: سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران.

سلطانی، پوری؛ فانی، کامران؛ زهادی، فیروزان (۱۳۸۵). سرعنوان‌های موضوعی فارسی ویراست سوم - پیوست دو. تهران: سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران.

سلطانی، پوری؛ فانی، کامران؛ مهناز رهبری اصل (۱۳۷۳). سرعنوان‌های موضوعی فارسی (ویرایش۲). تهران: سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران.

طباطبایی، علاءالدین (۱۳۸۸). ساختمان واژه و مقوله دستوری: تشخیص مقوله دستوری واژه‌ها، براساس ملاک‌های صرفی. تهران: پژوهشگاه فرهنگ، هنر و ارتباطات.

طبیب، مهدیس (۱۳۸۰). بررسی تطبیقی ساختار و روابط میان اصطلاحات علم کتابداری و اطلاع رسانی در سه اصطلاحنامه یونسکو، اریک و اسیس. فصلنامه کتاب، ۱۲(۳)، ص ۸۲- ۹۵.

فالک، جولیا (۱۳۷۱). زبان‌شناسی و زبان: بررسی مفاهیم بنیادی زبانشناسی. (خسرو غلامعلی‌زاده، مترجم). مشهد: آستان قدس رضوی، معاونت فرهنگی. (نشر اثر اصلی ۱۹۷۸)

فتاحی، رحمت‌الله ؛ ارسطوپور، شعله (۱۳۸۶). بررسی میزان همخوانی سرعنوان های موضوعی فارسی با کلید واژه های عنوان و فهرست مندرجات کتابهای فارسی در حوزه های علوم انسانی، علوم اجتماعی، علوم کاربردی و علوم محض. کتابداری و اطلاع رسانی، ۱۰(۳۹)، ص ۲۷-۳۹. بازیابی شده در ۸۸/۱/۲از http://www.aqlibrary.org/index.php?module=TWArticles&file=index&func=view_pubarticles&did=153&pid=10

فتوحی، محمود، عباسی، حبیب‌الله (۱۳۸۴). فارسی عمومی: درسنامه دانشگاهی. تهران: سخن.

کازرانی، مریم؛ اسدی گرگانی، فاطمه؛ رهادوست، فاطمه (۱۳۸۱). ارزیابی اصطلاحنامه‌های فارسی با استانداردهای ایزو: گامی به سوی بهره وری در ذخیره و بازیابی اطلاعات .کتاب ماه کلیات، ۶۲، ص ۸۹-۹۲.

کوهن، تامس (۱۳۸۳). ساختار انقلاب‌های علمی. (عباس طاهری، مترجم). تهران: قصه (تاریخ نشر اثر اصلی ۱۹۶۲).

مدرسی، فاطمه (۱۳۸۵). از واج تا جمله: فرهنگ زبان‌شناسی — دستور. تهران: چاپار.

میرزابیگی، مهدیه (۱۳۸۴). بررسی روابط میان اصطلاحات مشترک در اصطلاحنامه فرهنگی فارسی: اصفا و اصطلاحنامه بین المللی توسعه فرهنگی یونسکو. دفتر کتابداری، ۳۹(۴۳)، ص ۲۳۳-۲۴۶.

میرزابیگی، مهدیه؛ حری، عباس (۱۳۸۵). بررسی تأثیرپذیری اصطلاحنامه از ورود اصطلاحات جدید با رویکردی بر اصطلاحنامه اریک. دفترکتابداری ۴۰ (۴۵)، ص ۱۲۵-۱۴۶

Akhondi, S. Horri, A. (2006). Structural analysis of database Persian descriptors of Iranian Information and Documentation Center. *Library science*, 40 (45), 37-62. (In Persian)

AmirHosseini, M. (2005). Quantitative evaluation of network connectivity across descriptors of religious jurisprudence, logic, Islamic philosophy, theology and Quran thesaurus. Paper presented in National conference of Thesaurus and its application in digital environment, Qom. (In Persian)

AmirHosseini, M. (2005).Quantitative assessment of the terminological relationships in terms of CAB agricultural thesaurus. *national studies on librarianship and information organization*, 16 (4), 13-32. (in Persian)

AmirHosseini, M. (2007a).Qualitative and quantitative evaluation of effective factors in information storage and retrieval in Persian thesaurus. *Ketab-e Mah-e Kolliyat*, 118, 100-104. (In Persian)

AmirHosseini, M. (2007b).Quantitative assessment of the balance of the Islamic sciences thesaurus, move towards analysis of empowerment of vocabulary entry. *Ketab-e Mah-e Kolliyat*, 118, 16-29. (In Persian)

AmirHosseini, M. (2007c).Quality evaluation of Asfa and chemistry thesaurus based on the views of content, structure, physical and thesaurus management. *Ketab-e Mah-e Kolliyat*, 118, 65-72. (In Persian)

Bozorgi, A. (1991). Assessment of the compatibility of titles and topics in Farsi books. *national studies on librarianship and information organization*, 2 (1-4), 114-145. (In Persian)

Chaplan, M. (1995). Mapping "Laborline Thesaurus" Terms to Library of Congress Subject Headings: Implications for Vocabulary Switching. *Library Quarterly*. 65(1), 39-61.

Chen, Y., Gu, H., Perl, Y., & Geller, J. (2009 a). Structural group-based auditing of missing hierarchical relationships in UMLS. *Journal of Biomedical Informatics, 42*(3), 452-467.

Chen, Y., Gu, H., Perl, Y., Geller, J., & Halper, M. (2009 b). Structural group auditing of a UMLS semantic type's extent. *Journal of Biomedical Informatics, 42*(1), 41-52.

Ed, S., Antoine, I., Clay, R., & Dan, K. (2008). LCSH, SKOS and linked data, *Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications* (pp. 25-33). Berlin, Germany: Dublin Core Metadata Initiative.

Falk, J. (1978). *Linguistics and Language.* (Khosro Gholamalizadeh, Translator). Mashhad: Astan-e Ghods-e Razavi. (In Persian)

Fattahi, R. Arastoopoor, S. (2007) Correlation between Persian subject headings and table of contents as a keyword Persian Persian books in the areas of humanities, social sciences, applied sciences and pure sciences. *Library and information science,* 10 (39), 27-39. (In Persian)

Fernandez, C.W. (1991). Semantic Relationships between Title Phrases and LCSH. *Cataloging & Classification Quarterly*, 13(1), 51 – 77.

Fotuhi, M.; Abbasi, H.A. (2005). *Persian Language: university textbook*. Tehran: Sokhan. (In Persian)

Frank, E., & Paynter, G. W. (2004). Predicting Library of Congress Classifications from Library of Congress Subject Headings. *Journal of the American Society for Information Science & Technology*, *55*, 214-227.

Harper, C. A. (2006). Encoding library of congress subject headings in SKOS: authority control for the semantic web, *Proceedings of the 2006 international conference on Dublin Core and Metadata Applications: metadata for knowledge and learning* (pp. 89-94). Manzanillo, Colima, Mexico: Dublin Core Metadata Initiative.

Kai, E., Heiner, S., & Magnus, P. (2007). Interactive thesaurus assessment for automatic document annotation, *Proceedings of the 4th international conference on Knowledge capture* (pp. 103-110). Whistler, BC, Canada: ACM.

Kazerani, M. Asadi Ghorghani, F. Rahadoost, F. (2002). Evaluation of Persian thesauri with ISO standards: a step towards efficiency in information storage and retrieval. Ketab e Mah, 62. 89-92. (In Persian)

Khani S., Alibeyg M.R., Haghani, H. (2007). Structures of Pubmed and EMbase databases with NISO standard of the thesauri: to assessment of these databases' indexing methods. *Journal of Health Administration*. 10(27) ; 27 – 32. (In Persian)

Khojoui, F. (1993). The strengths and weaknesses of the hierarchical relationships in ASM and SDIM metallurgy thesaurus. *Payam-e Ketbkhaneh*, 3 (2,3), 22-32. (In Persian)

Kuhn, Thomas (1962). *The Structure of Scientific Revolutions*. (Abbas Thari, Translator). Tehran: Ghesseh. (In Persian)

Mirzabeigi, M (2005). Relationship between common terms in Persian Cultural Thesaurus: Asfa and UNESCO. *Library science*, 39(43), 233-246. (In Persian)

Mirzabeigi, M; Horri, A. (2006). The effect of the entrance of new terms in ERIC thesaurus. *Library science*, 40 (45), 125-146. (In Persian)

Modarresi, F. (206). *A dictionary of linguistics and phonetics.* Tehran: Chapar. (In Persian)

Okhovati, M ., DAvarpanah, M ., Dayani, M ., & Shabani, B . (2009). Analysing and comparing the concept network of Persian Subject Headings with the concept network of Persian LIS literature based on domain analysis. *Foundations of Education,10*(2).Retrievedfrom http://fedu.um.ac.ir/index.php/education/article/view/1594. (In Persian)

Sarmad, Z., Bazarghan, A., Hejazi, E. (1998). *Research Methods in the Behavioral Sciences*. Tehran: Aghah. (In Persian)

Soltani, P.; Fani, K.; Rahbari Asl, M. (1994). *Persian subject headings* (2ed edition). Tehran: National Library and Archives Organisation of Iran. (In Persian)

Soltani, P.; Fani, K.; Rahbari Asl, M. (2002). *Persian subject headings* (3th edition). Tehran: National Library and Archives Organisation of Iran. (In Persian)

Soltani, P.; Fani, K.; Rahbari Asl, M. (2003). *Persian subject headings* (3th edition, first appendix). Tehran: National Library and Archives Organisation of Iran. (In Persian)

Soltani, P.; Fani, K.; Rahbari Asl, M. (2006). *Persian subject headings* (3th edition, 2ed appendix). Tehran: National Library and Archives Organisation of Iran. (In Persian)

Steve, O. (2008). Bibliographic Supernova. *Library Journal, 133*(3), 132.

Tabatabaei, A. (2009). *Vocabulary Structure and grammatical category: identifying grammatical category of words, based on morphological criteria*. Tehran: Institute of Culture, Arts and Communications. (In Persian)

Tabib, M. (2001). A comparative study of the structure and relationships between terms of LIS in the UNESCO, Eric and Assisi Thesaurus, *national studies on librarianship and information organization*, 12(3). 82-95. (In Persian)

Yi, K. (2010). A semantic similarity approach to predicting Library of Congress subject headings for social tags. *Journal of the American Society for Information Science and Technology, 61*(8), 1658-1672.

Yi, K., & Chan, L. (2010). Revisiting the syntactical and structural analysis of Library of Congress Subject Headings for the digital environment. *Journal of the American Society for Information Science and Technology, 61*, 677-687.

Yi, K., & Chan, L. M. (2009). Linking folksonomy to Library of Congress subject headings: an exploratory study. *Journal of Documentation, 65*(6), 872 – 900.

Yi, K., & Chan, L.M. (2008). A visualization software tool for Library of Congress subject headings. Proceedings of the 10th International Conference of International Society for Knowledge Organization. In C. Arsenault & J.T. Tennis (Eds.), Advances in knowledge organization (Vol. 11, pp. 170–176).Würzburg, Germany: Ergon.

Youngok, C., Ingrid, H.-Y., & Bill, K. (2007). Retrieval effectiveness of table of contents and subject headings, *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries* (pp. 103-104). Vancouver, BC, Canada: ACM.Ed, S., Antoine, I., Clay, R., & Dan, K. (2008). LCSH, SKOS and linked data, *Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications* (pp. 25-33). Berlin, Germany: Dublin Core Metadata Initiative.

Zhoadi, F (2010). A comparative study of the form and content of the Subject headings of Islam in the Persian Subject Headings Authority database with that of the Library of Congress. *Library and information science,* 4 (12) at http://www.aqlibrary.org/index.php?module=TWArticles&file=index&func=view_pubarticles&did=743&pid=10. (In Persian)